Exhibit B

Volume 12 Number 22 1984 Nucleic Acids Research

Comparison of the whey acidic protein genes of the rat and mouse

S.M.Campbell and J.M.Rosen
Department of Cell Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA, and

L.G.Hennighausen+, U.Strech-Jurk* and A.E.Sippel
Institut für Genetik, Universität zu Köln, 5000 Köln 41, FRG

Received 6 August 1984; Revised and Accepted 22 October 1984

ABSTRACT

Whey acidic protein (WAP), a hormonally-regulated 14,000 dalton cysteine-rich protein, is the principal whey protein found in rodent milk. Genomic clones encompassing both the 2.8 Kb rat and 3.3 Kb mouse WAP genes have been characterized. The genes consist of four exons and three introns. The middle two exons encode the two cysteine-rich regions which probably form separate protein domains. Homology in the 5' flanking DNA of the mouse and rat extends at least 325 bp upstream of the putative CAP site, including a precisely conserved stretch of 50 bp around the unusual TATA and CAAT sites. The homology previously observed between the 3' noncoding sequences of the rat and mouse mRNAs extends at least 20 bp into the 3' flanking region. Several potential glucocorticoid receptor binding sites have been found in the 5' flanking region of the WAP gene. The conservation of the 5' flanking region of the WAP genes may be related to regulation of expression of WAP by peptide and/or steroid hormones.

INTRODUCTION

During lactation the mammary gland secretes large amounts of milk protein which consists predominantly of the acid-precipitable caseins and several whey proteins (1). WAP has so far been discovered only in the milk of rats, mice and rabbits (2; J.C. Mercier and L.G. Hennighausen, unpublished results), and not in the better-characterized ruminant milks. In the milk of the lactating rodent WAP is several-fold more abundant than α-lactalbumin (3). In the lactating rat mammary gland, WAP mRNA comprises 15% of the total poly(A)$^+$ mRNA (4,5). Studies using rat mammary gland explant cultures have shown that the hormonal induction of WAP is responsive to both glucocorticoids and prolactin, although the steroid hormone appears to be of primary importance (5).

Cloned WAP cDNA's from mouse and rat have been sequenced and the respective protein sequences deduced. (2,6,7) The mouse protein contains 134 amino acid residues while the rat protein contains 137. The most striking feature of these proteins is the large number of cysteine residues arranged into two domains with similar arrangements of the six cysteines in each. The signal peptide and cysteine domain 1 are more conserved at the amino acid level than

domain 2. Additionally, the 3' noncoding region is better conserved than the coding portions of the mRNA.

Although the function of WAP is unknown, certain features of its primary structure suggest comparisons with other proteins of known function. The arrangement of cysteines in each of the two domains of WAP is similar to that seen in the other "Four Disulfide Core Proteins" (7-9). These proteins, which include wheat germ agglutinin, several snake venom toxins and the neurophysins, share a common pattern of cysteines which in some cases have been shown to form specific intramolecular disulfide bonds and folding patterns (8,9). Gilbert (10) and Blake (11) have suggested that exons of eukaryotic genes may encode structural or functional domains of the protein. Exon multiplication seems to have played a role in the evolution of the collagen (12), ovomucoid (13) and α-fetoprotein (14) genes. The similar patterns of the cysteine clusters in the two WAP domains suggest that analogous mechanisms have operated in the case of WAP, with both domains evolving from a single primordial domain by means of an intragenic duplication (2).

Genomic clones from both the rat and mouse WAP genes have been isolated and used to derive the sequences of the coding exons, intron/exon splice junctions, and of the 5' and 3' flanking sequences. The exon structure of the WAP genes reflects the structure of the proteins. In addition, homology in the 5' and 3' flanking DNA of the mouse and rat WAP genes probably reflects conservation of sequences important to the expression and/or processing of the WAP mRNA.

MATERIALS AND METHODS

Materials

All restriction endonucleases were purchased from Boehringer Mannheim, New England Biolabs, Amersham, or Bethesda Research Laboratories. DNA polymerase I, DNA polymerase I Klenow fragment, T4 DNA ligase, and T4 polynucleotide kinase were from Bethesda Research Laboratories. Calf intestinal phosphatase was from Boehringer Mannheim. NACS-Prepak columns were obtained from Bethesda Research Laboratories.

Library Screening

A genomic DNA library of the GR mouse (15) was screened by *in situ* plaque hybridization (16). Fifty petri dishes each containing ~$2x10^4$ phage plaques were screened with the cloned insert of pWAP1 (1) as a hybridization probe. The hybridization probe was isolated by preparative gel electrophoresis and was radioactively labeled to a specific activity of ~10^8 cpm/μg DNA by nick

translation (17). A Charon 4A library of rat DNA generated by partial Hae III digestion was screened as described previously (18).

Gel Transfer and Hybridization

Total BALB/c mouse liver DNA or recombinant mouse DNA was digested with restriction endonucleases, electrophoresed on agarose gels, and transferred to nitrocellulose filters as described (19). The filters were hybridized to nick-translated pWAP6 DNA in order to detect the mouse WAP gene (1). The filters were washed in 0.2 X SSC at 65°C, and the hybridization signal was visualized by autoradiography. Transfers and hybridization of rat DNA were done as described (18).

Sandwich Hybridization

To visualize 5' regions not represented in the rat cDNA clone (2), an RNA sandwich hybridization technique was employed. The DNA of interest was digested with the appropriate restriction enzymes and electrophoresed on agarose gels. Bidirectional transfers were done using the method of Smith and Summers (20), and the filters baked for 4 hrs. The filters were prehybridized in 5 ml of 50% formamide, 0.6 M NaCl, 5 mM EDTA, 50 mM Tris-HCl pH 7.4, 0.1% SDS, 0.04% BSA, 0.04% polyvinylpyrrolidine, 0.04% Ficoll®, and 300 µg/ml sheared salmon sperm DNA (hybridization solution). After 4 hrs at 37°C, 5 µg of lactating mammary gland poly(A)$^+$ RNA was added and incubated at 37°C for 12 hrs. The filters were washed for 1 hr at 68°C in 2 X SSC, 0.1% SDS, then placed in a seal-a-bag containing 5 ml hybridization solution. Denatured nick-translated cDNA probe was added and the filters hybridized at 37°C overnight. The filters were washed in 2 X SSC, 0.1% SDS, first at room temperature for 10 min and then at 68°C for 1 hr. The hybridization patterns were visualized by autoradiography.

Plasmid Subcloning

For the mouse gene, the 7.2 Kb Eco RI fragment of λWAP24 was subcloned into pBR322 and into pUR250 (21). For the rat gene, the 1.1 and 2.4 Kb Eco RI fragments of WAPλ1 and the 3.3 Kb fragment of WAPλ2 were subcloned into pBR325. For dideoxy sequencing, these subclones were digested with Hae III and subcloned into the Sma I site, or with Sau 3A and subcloned into the Bam HI site of M13mp8. The digested WAP subclone DNA was treated with calf intestinal phosphatase before ligation to prevent multiple insertions. Single stranded DNA was isolated from the subclones and screened by hybridization with nick-translated WAP cDNA or by sandwich hybridization to isolate clones containing WAP exon sequences.

DNA Sequencing

The mouse sequence was obtained by 3' end labeling with (α-^{32}P) dNTP and Klenow fragment of DNA polymerase I (22). Fragments were separated on agarose gels, electroeluted and sequenced according to Maxam and Gilbert (23). Fragments subcloned in pUR250 were sequenced directly without purifying fragments (21). Rat WAP sequence was obtained by labeling either at 3' ends with (α-^{32}P) dNTP and Klenow (22), or at 5' ends with (γ-^{32}P) ATP and polynucleotide kinase (23). The end-labeled fragments were then digested with a second enzyme to generate single-end-labeled fragments, resolved by electrophoresis in low melting point agarose, and purified by elution from NACS-Prepak columns. Sequencing was then carried out by the procedure of Maxam and Gilbert (23). DNA was routinely precipitated from 2.5 M NH_4CH_3COO with 2.5 volumes of ethanol, and was desalted by precipitation from 26 µl of 300 mM NH_4CH_3COO acetate with 1 ml of ethanol, followed by lyophilization from 20 µl of distilled water. Dideoxy sequencing using M13mp8 was done by the method of Smith (24). Computer analysis of the sequence data was done using the HELIX Sequence Information System (25).

RESULTS AND DISCUSSION

Cloning of the WAP Genes, Identification of Coding Regions, and Sequencing

Two genomic clones, WAPλ1 and WAPλ2 (Fig. 1A), were isolated from the rat DNA library. WAPλ1 contains a 1.1 Kb Eco RI fragment, and WAPλ2 contains a 3.3 Kb Eco RI fragment, which hybridized to a 435 bp WAP cDNA clone. Both the 1.1 and 3.3 Kb fragments contain a unique Hinc II site, also present in the cDNA clone (1). This site is identical in both the 1.1 and 3.3 Kb fragments by DNA sequencing (data not shown). A probe consisting of the portion of the WAP cDNA extending from the unique Taq I site to the 3' end hybridized only to the 3.3 Kb band. This indicated that the 1.1 Kb piece lacks the 3'-most portions of the coding region. Since the restriction maps of WAPλ1 and WAPλ2 do not overlap outside of the 1.1 and 3.3 Kb pieces (Fig. 1A), it was concluded that WAPλ1 contains the 5' end and 5' flanking DNA, while WAPλ2 contains the 3' end and 3' flanking DNA. From the map of WAPλ1, it was predicted that the 5'end of the gene would reside in either the 1.1 Kb piece or the neighboring 2.4 Kb Eco RI fragment. Sandwich blotting to localize the 5' end of the gene detects the 2.4 Kb Eco RI piece of WAPλ1; this fragment does not hybridize to the ^{32}P-labeled WAP cDNA alone. The gene thus lies on three Eco RI fragments within two phage clones.

A genomic clone, λWAP24 (Fig. 1B), isolated from the GR mouse library,

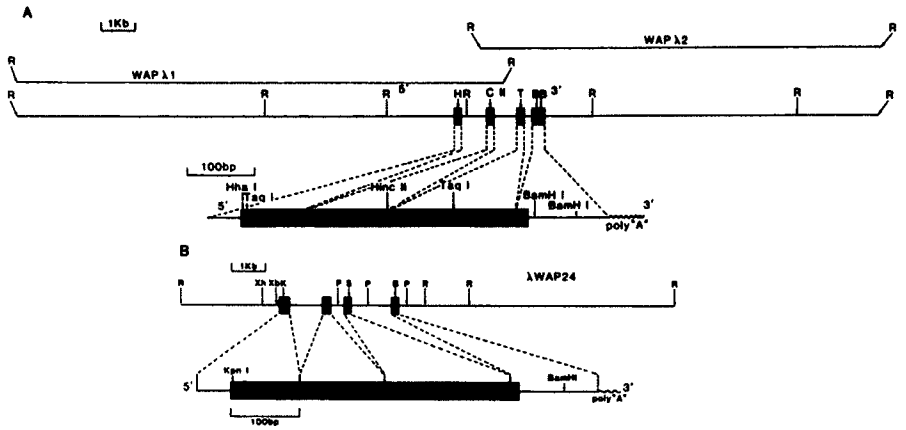


Figure 1 Structure of the Mouse and Rat WAP Genes. A) The Rat Whey Acidic Protein Gene. R = Eco RI, H = Hha I, C II = Hinc II, T = Taq I, B = Bam HI. Solid boxes in the genomic portion of the figure represent exons. The heavy line in the mRNA diagram represents the protein coding portion of the mRNA. B) The Mouse Whey Acidic Protein Gene. Xh = Xho I, Xb = Xba I, S = Sal I, P = Pst I.

contained a 7.2 Kb Eco RI fragment which hybridized to ^{32}P-labeled pWAP6, a plasmid with a 450 bp WAP specific cDNA insert. Southern blot analysis of total Eco RI-digested DNA also revealed the presence of a single 7.2 Kb fragment. Thus, the entire mouse WAP gene lies on this 7.2 Kb Eco RI fragment which carries a unique Bam HI restriction site, also present on the cloned cDNA (7; see Fig. 1B).

Partial restriction maps of the WAP clones and the strategy used for sequencing are shown in Fig. 2. The sequence was obtained from both strands of the DNA in all cases. The conclusions derived from the sequencing are discussed below. The complete sequence of both the rat and mouse genes is found in Fig. 3.

Exon Structure of the WAP Gene Mirrors the Structure of the Protein

The WAP gene extends over 3.3 Kb of the genome in the mouse, and 2.8 Kb in the rat, and is composed of four exons divided by three introns (Fig. 1A and B). The difference in the sizes of the genes is primarily due to the difference in the third introns - the size in the mouse is 1.10 Kb, in the rat this intron is ~500 bp. The first exon encodes 26 nucleotides of the 5' untranslated region of the mRNA in the mouse, 33 in the rat, plus the sequences encoding the 19 amino acid signal peptide and the first 10 amino acids of the mature protein (Fig. 3). The second exon is 135 bp long in both

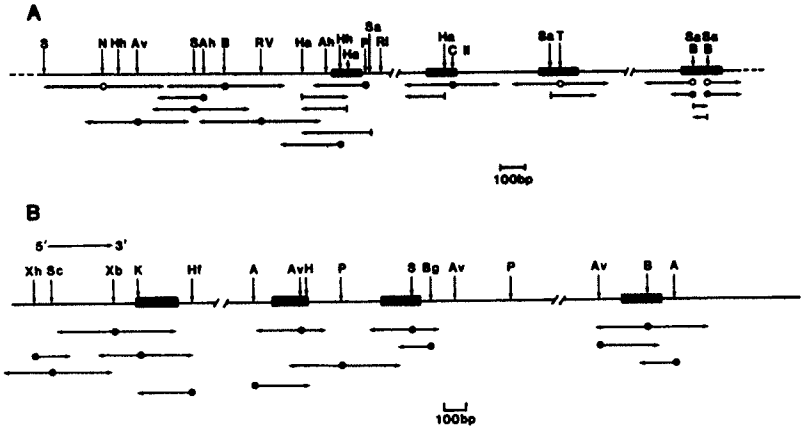
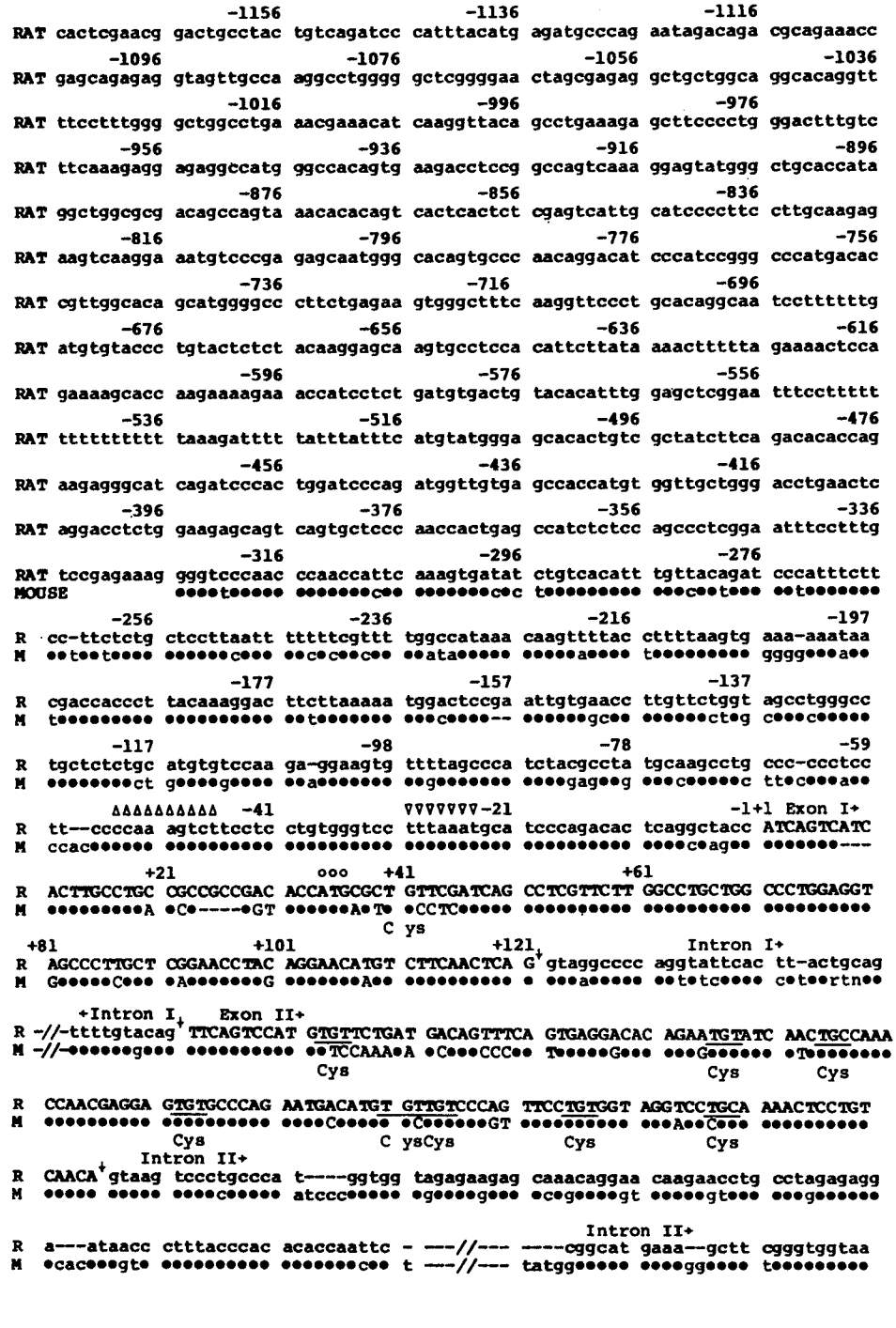


Figure 2 Sequencing Strategy for the WAP Genes. A) Rat Gene. Ah = Aha III, Av = Ava I, B = Bam HI, C II = Hinc II, Ha = Hae III, Hh = Hha I, N = Nco I, RI = Eco RI, RV = Eco RV, S = Sst I, Sa = Sau 3A, T = Taq I. Fragments were labeled on the 5' (solid circles), or 3' (open circles) ends, or were sequenced by dideoxy method (vertical line). Solid boxes on line indicate exons. B) Mouse Gene. A = Alu I, Av = Ava II, B = Bam HI, Bg = Bgl II, H = Hinc II, Hf = Hinf I, K = Kpn I, P = Pst I, S = Sal I, Sc = Sac I, Xb = Xba I, Xh = Xho I. Sequences were derived either from purified restriction fragments or by direct sequence analysis of fragments cloned into the polylinker of the vector pUR250 (4,22). Solid boxes indicate exons.

the mouse and rat, while the third exon comprises 165 bp in the mouse and 168 in the rat. It has been hypothesized that the second and third exons of the WAP gene arose via an intragenic duplication of a primordial gene containing a single cysteine domain (2). Each exon encodes precisely 1 cysteine domain; additionally, the 3' ends of the second and third exon are both located at analogous positions with respect to the cysteine clusters; the splice junction splits the fourteenth codon 3' to the double cysteine in each domain. The fourth exon consists of 146 nucleotides in the mouse and 161 in the rat, encoding the last 4 amino acids of the mouse protein and 8 of the rat, the stop codon, and 131 bp of the conserved 3' noncoding region in the mouse, 134 bp in the rat (Fig. 3). The sequence of the introns is only slightly more divergent than that of the exons.

Many protein molecules are composed of separate regions of folding known as protein or structural domains. Gilbert (10) and Blake (11) suggested that exons of eukaryotic genes might correspond with particular structural domains of defined function, providing organisms with the structural building blocks



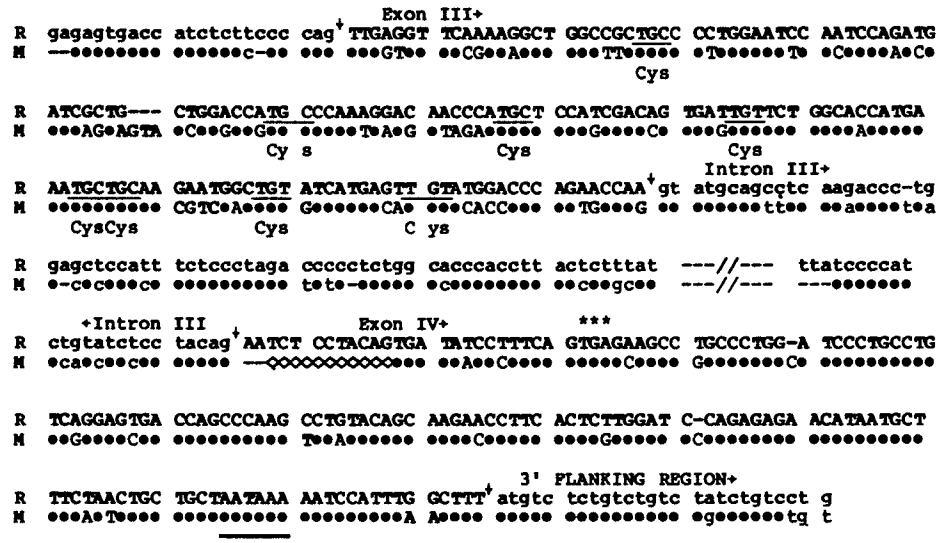


Figure 3 Sequence of the Rat and Mouse WAP Genes. The CAAT box (Δ), TATA box (∇), initiation codon (○), cryptic rat splice consensus region (◇), and termination codon (*) are indicated. Triplets encoding cysteines are underscored; the polyadenylation signal is doubly underscored. Splice junctions are indicated by arrows (+) as is the end of the mRNA.

from which novel protein functions would be made. For example, in the cases of collagen (12) and α-fetoprotein (14) there is evidence from amino acid sequence and genomic structure to establish that the genes evolved by intragenic amplification of exons. In the case of ovomucoid (13), exon amplification gave rise to three functional domains, connected by short peptides and with disulfide bonds formed strictly within the domains. And although the position of these cysteines is constant within each domain, there is only low homology of the amino acid sequences in each domain, exactly as is found in WAP (2). In another abundant milk protein gene, the β-casein gene of the rat, there is also evidence for exon amplification as a mechanism of evolution (26). WAP provides another possible example for this mechanism of the generation of protein diversity.

Splice Junctions of the WAP Exons

The splice junctions of the WAP gene shown in Fig. 4 are similar to the splice junction consensus sequences compiled by Mount (27), and also conform to Chambon's rules (28). An interesting feature is the 12 bp (4 amino acid) insertion found in the rat gene. This insertion involves a duplication (10 of

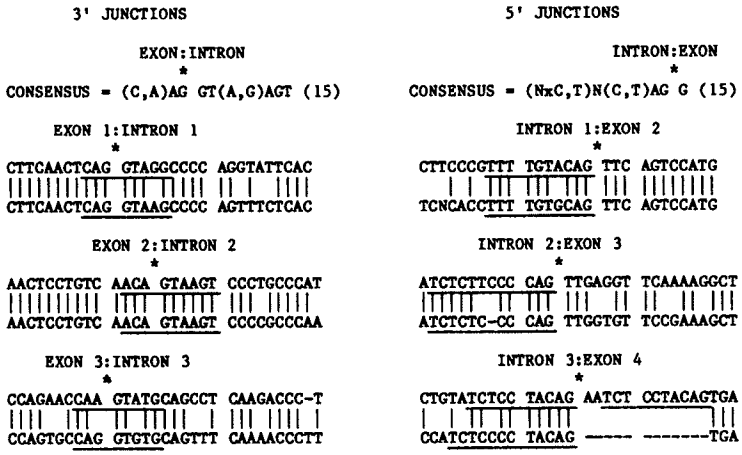


Figure 4 A Comparison of Rat and Mouse WAP Gene Splice Junctions. Splice junction consensus sequences are underlined.

12 bases homology) of the splice consensus region of the rat relative to the mouse, although it is impossible to determine whether this phenomenon arose via a duplication of this region in the rat or a deletion of one copy of the consensus in the mouse. Another possibility would involve activation of a cryptic splice junction sequence in the intron of the rat gene. The factors which cause the selection of the 5'-most splice consensus in the rat instead of the 3'-most consensus are unknown.

Conservation of the 3' Flanking DNA of the WAP Gene

The sequences of the rat and mouse WAP gene fourth exons and 3' flanking regions are shown in Fig. 3. As reported previously (1), there is greater homology between the rat and mouse genes in the 3' noncoding region of the gene (91%) than within the coding regions (82%). It is likely that some important function of this region has resulted in its unusual conservation. The homology extends 22 bases 3' to the putative message end. The 3' flanking DNA is repetitive in sequence, consisting of T residues in every other position, alternating mainly with G and C residues. This pattern ends after +22, at least in the rat. A computer search (25) of mammalian gene sequences in the Genbank™ sequence database did not reveal any other genes with similar stretches in analogous positions. There do not appear to be regions of homology with U4 snRNA, as have recently been reported surrounding polyadenylation sites (29). We also do not observe any sequence closely resembling that found by Benoist *et al.* in 5 of 11 mRNA's, although the region surrounding the

polyadenylation site is AT-rich as seen in those mRNAs (30).

Comparison of the Rat and Mouse 5' Flanking Regions

The sequences of the rat and mouse 5' flanking DNA and first exons are compared in Fig. 3. The 5' end of the mouse mRNA has been tentatively localized by *in vitro* transcription of the WAP gene using a whole HeLa cell transcription extract (L.G. Hennighausen and M. Theisen, unpublished results), and by sequencing full-length cDNA clones (31). The CAP site is probably at the A residue 26 nucleotides upstream from the initiation site of translation in the mouse, 29 in the rat. The homology between the mouse and rat sequences is 83%, and extends at least to position -325. Further data indicate that the homology extends as far as -800 (unpublished observations). An unusual form of the Goldberg-Hogness 'TATA' box (28), TTTAAAT, is present in both the rat and mouse genes at a position 30 bases upstream from the predicted initiation site. A sequence similar to the 'CAAT' box (28) sequence, CAAAGTCT, is located around -55. A region of 50 bases including these two postulated regulatory sites is perfectly conserved between the mouse and rat genes. In contrast, the 5' flanking regions of the rat β-casein (S.M. Campbell, W.K. Jones, and J.M. Rosen, unpublished observation) and α-lactalbumin genes (32), whose expression are also regulated by prolactin and glucocorticoids, demonstrate no such homology with the rat WAP gene.

The rat and mouse diverged ~17 million years ago (33). Most quoted rates of evolution of silent codon positions or of introns are in the range of 3.5-5.5 x 10^{-9} mutations/site/year. These values indicate that the overall homology found in the WAP 5' flanking DNA is unlikely to be significant. However, there are local regions of higher homology, particularly the precisely conserved 50 bp surrounding the 'TATA' and 'CAAT' boxes. These may represent sequences of functional importance conserved through evolution.

The unusual 'TATA box' we find in the WAP gene (TTTAAAT) is also found in the γ- (18) and α-casein (34) genes, but not in the β-casein gene (26). A computer search of 105 eukaryotic 5' flanking sequences in Genbank™ found only 3 genes with a 'T' in the second position of their 'TATA boxes'. This was found in the rat preprocarboxypeptidase gene, whose 'TATA box' sequence is TTTAAAA (35), the bovine opsin gene (TTCATAA) (36), and this rat preprosomatostatin gene (TTTAAAA) (37). Thus the TTTAAAT found in the genes mentioned above appears to be a relatively specific feature of a small number of genes encoding major milk proteins. Preliminary *in vitro* transcription analysis of the α- and γ-casein genes indicates that the TTTAAAT sequence is a weak promoter in this system (38). The transition of an A to a C, T, or G in the second position of the TATA sequence has been shown to reduce the *in vitro*

efficiency of the promoter sequence in other genes (35). Of interest is the lack of this unusual element in the β-casein gene, which is expressed to a greater extent than the other three milk protein genes during lactation. The role that the altered TATA sequence may play in tissue and differentiation specific regulation of mammary gland gene expression remains to be established.

The mechanism by which WAP gene expression is regulated by glucocorticoids is unknown. One thousand one hundred ninety seven bp of 5' flanking DNA of the rat gene have been sequenced and within this sequence several short elements related to the family of octanucleotides reported by Payvar et al. (40) in sequences shown by DNase I footprinting to bind the glucocorticoid receptor have been observed. These sequences are at -394, -598, -626 and -1127 (Fig. 3). In addition, a 9 bp subset of the putative progesterone receptor binding sequence of Mulvihill et al. (41) is present in both the mouse and rat at -277 (rat coordinates, Fig.3). A competitive DNA cellulose binding assay (42) has identified multiple glucocorticoid receptor binding sites between -566 and -1250 in the rat WAP gene, some of which have affinities comparable to those observed for the MMTV LTR DNA (M. Pfahl, unpublished results). Studies are underway to further characterize the functional importance of these regions and to define other hormone-responsive elements in the WAP gene. The structural information reported in this manuscript is a necessary prerequisite for the definition of these sequences by both direct hormone receptor binding studies including DNAse I footprinting and DNA-mediated gene transfer experiments.

ACKNOWLEDGEMENTS

L.G. Hennighausen thanks H. Ponta (Karlsruhe) for the mouse DNA library, R. Lange (Köln) and U. Borgmeyer (Köln) who helped with subcloning, W. Vielmetter (Köln) and P. Leder (Boston) for providing laboratory space, and J. Battey (Boston) who helped with computer studies. This work was supported by the Deutsche Forschungsgemeinschaft through the SBF74/E5.

S.M. Campbell thanks L. Jagodzinski, T. Sargent and J. Bonner for the rat DNA library, C. Green (Liverpool) for help in the screening of the library, and C. Lawrence, for providing computer facilities. Part of this work was supported by NIH grant CA16303, Welch Foundation grant G-969, and an NIH Medical Scientist Training Program award (S.M. Campbell).

+Present address: Department of Genetics, Harvard Medical School, 45 Shattuck Street, Boston, MA 02115, USA

*Present address: Zentrum für Molecular Biologie Heidelberg, Universität Heidelberg, 69 Heidelberg, FRG

REFERENCES

1. Hennighausen, L.G. and Sippel, A.E. (1982) Eur. J. Biochem. 125, 131-141.
2. Hennighausen, L.G., Sippel, A.E., Hobbs, A.A. and Rosen, J.M. (1982) Nucl. Acids Res. 10, 3733-3744.
3. Zamierowski, M.M. and Ebner, K.E. (1980) J. Immun. Methods 36, 211-220.
4. Richards, D.A., Rodgers, J.R., Supowit, S.C. and Rosen, J.M. (1981) J. Biol. Chem. 256, 526-532.
5. Hobbs, A.A., Richards, D.A., Kessler, D.T. and Rosen, J.M. (1982) J. Biol. Chem. 257, 3598-3605.
6. Dandekar, A.M., Robinson, E.A., Appella, E. and Qasba, P.K. (1982) Proc. Natl. Acad. Sci. USA 79, 3987-3991.
7. Hennighausen, L.G. and Sippel, A.E. (1982) Nucl. Acids Res. 10, 2677-2684.
8. Drenth, J., Low, B.W., Richardson, J.S. and Wright, C.S. (1980) J. Biol. Chem. 255, 2652-2655.
9. Drenth, J. (1981) J. Biol. Chem. 256, 2601-2602.
10. Gilbert, W. (1978) Nature 271, 501.
11. Blake, C.C.F. (1978) Nature 273, 267.
12. de Crombrugge, B. and Paston, J. (1982) Trends Biochem. Sci. 7, 11-13.
13. Stein, J.P., Catterall, J.F., Kristo, P., Means, A.E. and O'Malley, B.W. (1980) Cell 21, 681-687.
14. Eiferman, F.A., Young, P.R., Scott, R.W. and Tilghman, S.M. (1981) Nature 294, 713-718.
15. Herrlich, P., Hynes, N.E., Ponta, H., Rahmsdorf, U., Kennedy, N. and Groner, B. (1981) Nucl. Acids Res. 9, 4981-4995.
16. Benton, W.D. and Davies, R.W. (1977) Science 196, 180-182.
17. Rigby, P.W.J., Dieckmann, M., Rhoades, C. and Berg, P. (1977) J. Mol. Biol. 113, 237-251.
18. Yu-Lee, L.Y. and Rosen, J.M. (1983) J. Biol. Chem. 258, 10794-10804.
19. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
20. Smith, G.E. and Summers, M.D. (1980) Anal. Biochem. 109, 123-129.
21. Ruther, U. (1982) Nucl. Acids Res. 10, 5765-5772.
22. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning - A Laboratory Manual, p. 113-116, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. Maxam, A.M. and Gilbert, W. (1980) Meth. Enzymol. 65, 499-560.
24. Smith, A.J.H. (1982) Meth. Enzymol. 65, 560-580.
25. Lawrence, C.B. (1984) Software Tools for Genetic Engineering and Nucleic Acid Sequence Analysis. Submitted.
26. Jones, W.K., Yu-Lee, L.Y., Clift, S.M., Brown T.L. and Rosen, J.M. (1984) In preparation.
27. Mount, S.M. (1982) Nucl. Acids Res. 10, 459-472.
28. Breathnach, R. and Chambon, P.A. (1981) Ann. Rev. Biochem. 50, 349-383.
29. Berget, S. (1984) Nature 309, 179-182
30. Benoist, C., O'Hare, K., Breathnach, R. and Chambon, P. (1980) Nucl. Acids Res. 8, 127-142.
31. Motojima, K. and Oka, T. (1983) Biochem. Biophys. Res. Commun. 116, 167-172.
32. Qasba, P. and Safaya, S. (1984) Nature 308, 377-380.
33. Miyata, T., Hayashida, H., Kikuno, K., Hasegawa, M., Kobayashi, M. and Koike, K. (1982) J. Mol. Evol. 19, 28-35.
34. Rosen, J.M., Jones, W.K., Campbell, S.M., Bisbee, C.H. and Yu-Lee, L.Y. (1984) Proceedings of the UCLA Symposium on Membrane Receptors and

Cellular Regulation, Kahn, C.R. and Czech, M.P. Eds., Allen R. Liss Inc., New York (in press).
35. Quinto, C., Quiroga, M., Swain, W., Nikovits, W., Standring, D., Pictet, R., Valenzuela, P. and Rutter, W. (1982) Proc. Natl. Acad. Sci. USA 79, 31-35.
36. Hogness, D. and Nathans, J. (1983) Cell 34, 807-814.
37. Montminy, M., Goodman, R., Horovitch, S. and Habener, J. (1984) Proc. Natl. Acad. Sci. USA 81, 3337-3340.
38. Yu-Lee, L.Y., Tsai, S.Y. and Rosen, J.M. (unpublished observations).
39. Compton, J.G., Schrader, W.T. and O'Malley, B.W. (1984) Submitted.
40. Payvar, F., DeFranco, D., Firestone, G., Edgar, B., Wrange, O., Okret, S., Gustaffson, J. and Yamamoto, K. (1983) Cell 35, 381-392.
41. Mulvihill, E.R., LePennec, J.P. and Chambon, P. (1982) Cell 24, 621-632.
42. Pfahl, M. (1982) Cell 31, 475-482.